Chapter 3

The Laplacian and Graph Drawing

3.1 The Laplacian Matrix

We being this section by establishing the equivalence of multiple expressions for the Laplacian.

The Laplacian Matrix of a weighted graph $G = (V, E, w)$, $w : E \to \mathbb{R}^+$, is designed to capture the Laplacian quadratic form:

$$\boldsymbol{x}^T \boldsymbol{L}_G \boldsymbol{x} = \sum_{(a,b)\in E} w_{a,b} (\boldsymbol{x}(a) - \boldsymbol{x}(b))^2. \tag{3.1}$$

We will now use this quadratic form to derive the structure of the matrix. To begin, consider a graph with just two vertices and one edge of weight 1. Let's call it $G_{1,2}$. We have

$$\boldsymbol{x}^T \boldsymbol{L}_{G_{1,2}} \boldsymbol{x} = (\boldsymbol{x}(1) - \boldsymbol{x}(2))^2. \tag{3.2}$$

Consider the vector $\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2$, where $\boldsymbol{\delta}_a$ is the elementary unit vector with a 1 in coordinate a. We have

$$\boldsymbol{x}(1) - \boldsymbol{x}(2) = \boldsymbol{\delta}_1^T \boldsymbol{x} - \boldsymbol{\delta}_2^T \boldsymbol{x} = (\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2)^T \boldsymbol{x},$$

so

$$(\boldsymbol{x}(1) - \boldsymbol{x}(2))^2 = \left((\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2)^T \boldsymbol{x}\right)^2 = \boldsymbol{x}^T (\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2)(\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2)^T \boldsymbol{x} = \boldsymbol{x}^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \boldsymbol{x}.$$

Thus,

$$\boldsymbol{L}_{G_{1,2}} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Now, let $G_{a,b}$ be the graph with just one edge between a and b. It can have as many other vertices as you like. The Laplacian of $G_{a,b}$ can be written in the same way:

$$\boldsymbol{L}_{G_{a,b}} = (\boldsymbol{\delta}_a - \boldsymbol{\delta}_b)(\boldsymbol{\delta}_a - \boldsymbol{\delta}_b)^T.$$

This is the matrix that is zero except at the intersection of rows and columns indexed by a and b, where it looks looks like

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Summing the matrices for every edge, we obtain

$$\boldsymbol{L}_G = \sum_{(a,b)\in E} w_{a,b}(\boldsymbol{\delta}_a - \boldsymbol{\delta}_b)(\boldsymbol{\delta}_a - \boldsymbol{\delta}_b)^T = \sum_{(a,b)\in E} w_{a,b}\boldsymbol{L}_{G_{a,b}}.$$

You can check that this agrees with the definition of the Laplacian from Section 1.2.3:

$$\boldsymbol{L}_G = \boldsymbol{D}_G - \boldsymbol{A}_G,$$

where

$$\boldsymbol{D}_G(a,a) = \sum_b w_{a,b}.$$

This formula turns out to be useful when we view the Laplacian as an operator. For every vector $\boldsymbol{x}$ we have

$$(\boldsymbol{L}_G\boldsymbol{x})(a) = d(a)\boldsymbol{x}(a) - \sum_{(a,b)\in E} w_{a,b}\boldsymbol{x}(b) = \sum_{(a,b)\in E} w_{a,b}(\boldsymbol{x}(a) - \boldsymbol{x}(b)). \tag{3.3}$$

From (3.1), we see that if all entries of $\boldsymbol{x}$ are the same, then $\boldsymbol{x}^T\boldsymbol{L}\boldsymbol{x}$ equals zero. From (3.3), we can immediately see that $\boldsymbol{L}\mathbf{1} = \mathbf{0}$, so the constant vectors are eigenvectors of eigenvalue zero. If the graph is connected, these are the only eigenvectors of eigenvalue zero.

Lemma 3.1.1. *Let $G = (V, E)$ be a graph, and let $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues of its Laplacian matrix, $\boldsymbol{L}$. Then, $\lambda_2 > 0$ if and only if G is connected.*

Proof. We first show that $\lambda_2 = 0$ if G is disconnected. If G is disconnected, then it can be described as the union of two graphs, G_1 and G_2. After suitably reordering the vertices, we can write

$$\boldsymbol{L} = \begin{bmatrix} L_{G_1} & 0 \\ 0 & L_{G_2} \end{bmatrix}.$$

So, $\boldsymbol{L}$ has at least two orthogonal eigenvectors of eigenvalue zero:

$$\begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix}.$$

where we have partitioned the vectors as we did the matrix $\boldsymbol{L}$.

On the other hand, assume that G is connected and that $\boldsymbol{\psi}$ is an eigenvector of $\boldsymbol{L}$ of eigenvalue 0. As

$$\boldsymbol{L}\boldsymbol{\psi} = \mathbf{0},$$

we have

$$x^T\boldsymbol{L}x = \sum_{(a,b)\in E} (\boldsymbol{\psi}(a) - \boldsymbol{\psi}(b))^2 = 0.$$

Thus, for every pair of vertices (a, b) connected by an edge, we have $\boldsymbol{\psi}(a) = \boldsymbol{\psi}(b)$. As every pair of vertices a and b are connected by a path, we may inductively apply this fact to show that $\boldsymbol{\psi}(a) = \boldsymbol{\psi}(b)$ for all vertices a and b. Thus, $\boldsymbol{\psi}$ must be a constant vector. We conclude that the eigenspace of eigenvalue 0 has dimension 1. □

Of course, the same holds for weighted graphs.

3.2 Drawing with Laplacian Eigenvalues

The idea of drawing graphs using eigenvectors demonstrated in Section 1.5.1 was suggested by Hall [Hal70] in 1970.

To explain Hall's approach, we first consider the problem of drawing a graph on a line. That is, mapping each vertex to a real number. It isn't easy to see what a graph looks like when you do this, as all of the edges sit on top of one another. One can fix this either by drawing the edges of the graph as curves, or by wrapping the line around a circle.

Let $\boldsymbol{x} \in \mathbb{R}^V$ be the vector that describes the assignment of a real number to each vertex. We would like vertices that are neighbors to be close to one another. So, Hall suggested that we choose an $\boldsymbol{x}$ minimizing

$$\boldsymbol{x}^T \boldsymbol{L} \boldsymbol{x} = \sum_{(a,b) \in E} (\boldsymbol{x}(a) - \boldsymbol{x}(b))^2. \tag{3.4}$$

Unless we place restrictions on $\boldsymbol{x}$, the solution will be degenerate. For example, all of the vertices could map to 0. To avoid this, and to fix the scale of the embedding overall, we require

$$\sum_{a \in V} \boldsymbol{x}(a)^2 = \|\boldsymbol{x}\|^2 = 1. \tag{3.5}$$

Even with this restriction, another degenerate solution is possible: it could be that every vertex maps to $1/\sqrt{n}$. To prevent this from happening, we impose the additional restriction that

$$\sum_{a} \boldsymbol{x}(a) = \mathbf{1}^T \boldsymbol{x} = 0. \tag{3.6}$$

On its own, this restriction fixes the shift of the embedding along the line. When combined with (3.5), it guarantees that we get something interesting.

As $\mathbf{1}$ is the eigenvector of smallest eigenvalue of the Laplacian, Theorem 2.2.2 implies that a unit eigenvector of λ_2 minimizes $\boldsymbol{x}^T \boldsymbol{L} \boldsymbol{x}$ subject to (3.5) and (3.6).

Of course, we really want to draw a graph in two dimensions. So, we will assign two coordinates to each vertex given by $\boldsymbol{x}$ and $\boldsymbol{y}$. As opposed to minimizing (3.4), we will minimize the sum of the squares of the lengths of the edges in the embedding:

$$\sum_{(a,b) \in E} \left\| \begin{pmatrix} \boldsymbol{x}(a) \\ \boldsymbol{y}(a) \end{pmatrix} - \begin{pmatrix} \boldsymbol{x}(b) \\ \boldsymbol{y}(b) \end{pmatrix} \right\|^2.$$

This turns out not to be so different from minimizing (3.4), as it equals

$$\sum_{(a,b) \in E} (\boldsymbol{x}(a) - \boldsymbol{x}(b))^2 + (\boldsymbol{y}(a) - \boldsymbol{y}(b))^2 = \boldsymbol{x}^T \boldsymbol{L} \boldsymbol{x} + \boldsymbol{y}^T \boldsymbol{L} \boldsymbol{y}.$$

As before, we impose the scale conditions

$$\|\boldsymbol{x}\|^2 = 1 \quad \text{and} \quad \|\boldsymbol{y}\|^2 = 1,$$

and the centering constraints

$$\mathbf{1}^T \boldsymbol{x} = 0 \quad \text{and} \quad \mathbf{1}^T \boldsymbol{y} = 0.$$

However, this still leaves us with the degenerate solution $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{\psi}_2$. To ensure that the two coordinates are different, Hall introduced the restriction that $\boldsymbol{x}$ be orthogonal to $\boldsymbol{y}$. To embed a graph in k dimensions, we find k orthonormal vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ that are orthogonal to $\mathbf{1}$ and minimize $\sum_i \boldsymbol{x}_i^T \boldsymbol{L} \boldsymbol{x}_i$. A natural choice for these is $\boldsymbol{\psi}_2$ through $\boldsymbol{\psi}_{k+1}$, and this choice achieves objective function value $\sum_{i=2}^{k+1} \lambda_i$.

The following theorem says that this choice is optimal. It is a variant of [Fan49, Theorem 1].

Theorem 3.2.1. *Let $\boldsymbol{L}$ be a Laplacian matrix and let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ be orthonormal vectors that are all orthogonal to $\mathbf{1}$. Then*

$$\sum_{i=1}^{k} \boldsymbol{x}_i^T \boldsymbol{L} \boldsymbol{x}_i \geq \sum_{i=2}^{k+1} \lambda_i,$$

and this inequality is tight only when $\boldsymbol{x}_i^T \boldsymbol{\psi}_j = 0$ for all j such that $\lambda_j > \lambda_{k+1}$.

Proof. Without loss of generality, let $\boldsymbol{\psi}_1$ be a constant vector.

Let $\boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_n$ be vectors such that $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ is an orthonormal basis. We can find these by choosing $\boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_n$ to be an orthonormal basis of the space orthogonal to $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$. We now know that for all $1 \leq i \leq n$

$$\sum_{j=1}^{n} (\boldsymbol{\psi}_j^T \boldsymbol{x}_i)^2 = 1 \qquad \text{and} \qquad \sum_{j=1}^{n} (\boldsymbol{x}_j^T \boldsymbol{\psi}_i)^2 = 1.$$

That is, the matrix with i,j entry $(\boldsymbol{\psi}_j^T \boldsymbol{x}_i)^2$ is *doubly-stochastic*[1].

Since $\boldsymbol{\psi}_1^T \boldsymbol{x}_i = 0$, Lemma 2.1.1 implies

$$\begin{aligned}
\boldsymbol{x}_i^T \boldsymbol{L} \boldsymbol{x}_i &= \sum_{j=2}^{n} \lambda_j (\boldsymbol{\psi}_j^T \boldsymbol{x}_i)^2 \\
&= \lambda_{k+1} + \sum_{j=2}^{n} (\lambda_j - \lambda_{k+1})(\boldsymbol{\psi}_j^T \boldsymbol{x}_i)^2, \qquad \text{by } \sum_{j=2}^{n} (\boldsymbol{\psi}_j^T \boldsymbol{x}_i)^2 = 1 \\
&\geq \lambda_{k+1} + \sum_{j=2}^{k+1} (\lambda_j - \lambda_{k+1})(\boldsymbol{\psi}_j^T \boldsymbol{x}_i)^2,
\end{aligned}$$

as $\lambda_j \geq \lambda_{k+1}$ for $j > k+1$. This inequality is only tight when $(\boldsymbol{\psi}_j^T \boldsymbol{x}_i)^2 = 0$ for j such that $\lambda_j > \lambda_{k+1}$.

[1] This theorem is really about *majorization*, which is easily established through multiplication by a doubly-stochastic matrix.

Summing over i we obtain

$$\begin{aligned}\sum_{i=1}^{k} \boldsymbol{x}_i^T \boldsymbol{L} \boldsymbol{x}_i &\geq k\lambda_{k+1} + \sum_{j=2}^{k+1} (\lambda_j - \lambda_{k+1}) \sum_{i=1}^{k} (\boldsymbol{\psi}_j^T \boldsymbol{x}_i)^2 \\ &\geq k\lambda_{k+1} + \sum_{j=2}^{k+1} (\lambda_j - \lambda_{k+1}) \\ &= \sum_{j=2}^{k+1} \lambda_j,\end{aligned}$$

where the inequality follows from the facts that $\lambda_j - \lambda_{k+1} \leq 0$ and $\sum_{i=1}^{k} (\boldsymbol{\psi}_j^T \boldsymbol{x}_i)^2 \leq 1$. This inequality is tight under the same conditions as the previous one.

□

The beautiful pictures that we sometimes obtain from Hall's graph drawing should convince you that eigenvectors of the Laplacian should reveal a lot about the structure of graphs. But, it is worth pointing out that there are many graphs for which this approach does not produce nice images, and there are in fact graphs that can not be nicely drawn. Expander graphs are good examples of these.

Many other approaches to graph drawing borrow ideas from Hall's work: they try to minimize some function of the distances of the edges subject to some constraints that keep the vertices well separated. However, very few of these have compactly describable solutions, or even solutions that can provably be computed in polynomial time. The algorithms that implement them typically use a gradient based method to attempt to minimize the function of the distances subject to constraints. This means that relabeling the vertices could produce very different drawings! Thus, one must be careful before using these images to infer some truth about a graph.